PCT/GB 97 / 01087

09/171553

The Patent Office

PRIORITY DOCUMENT

The Patent Office
Cardiff Road
Newport
Gwent
NP9 1RH

RECEIVED 02 JUN 1997 PCT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation and Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

Dated 13 MAY 1997

Patents Form 1/77

Patents Act 1977
(Rule 16)

10 FEB 1997
LONDON

The Patent Office

10 FEB 1997

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

12FEB97 E253168-
P01/7700 25.00

The Patent Office
Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference — MGH/PC/MG/P07557GB

2. Patent application number *(The Patent Office will fill in this part)* — **9702668.6**

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

   Q-ONE BIOTECH LIMITED
   Todd Campus
   West of Scotland Science Park
   GLASGOW G20 OXA

   Patents ADP number *(if you know it)*

   If the applicant is a corporate body, give the country/state of its incorporation — 6602247001

4. Title of the invention — PORCINE RETROVIRUS

5. Name of your agent *(if you have one)*

   "Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

   CRUIKSHANK & FAIRWEATHER
   19 Royal Exchange Square
   GLASGOW G1 3AE

   Patents ADP number *(if you know it)* — 547002

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*
   a) any applicant named in part 3 is not an inventor, or
   b) there is an inventor who is not named as an applicant, or
   c) any named applicant is a corporate body.
   See note (d))

Patents Form 1/77

Porcine Retrovirus

The present invention relates inter alia to porcine retrovirus (PoEV) fragments, in particular polynucleotide fragments encoding at least one porcine retrovirus expression product, a recombinant vector comprising at least one polynucleotide fragment, use of PoEV polynucleotide fragments in the detection of native porcine retrovirus, a host cell containing at least one PoEV polynucleotide fragment or a recombinant vector comprising at least one PoEV polynucleotide fragment, PoEV polypeptides, antibodies immuno-reactive with PoEV polypeptides, pharmaceutical compositions comprising recombinant PoEV polypeptides for use as prophylactic and/or therapeutic agents and uses of PoEV polynucleotide fragments and/or polypeptides in medicine, including veterinary medicine and in the preparation of medicaments for use in medicine, including veterinary medicine.

Porcine retrovirus (PoEV) is an endogenous (genetically acquired) retrovirus isolated from pigs and expressed in cell lines derived from porcine material. There are no known pathogenic effects associated with the virus per se in its natural host although the virus appears to be associated with lymphomas in pigs and related viruses are associated with leukaemias and lymphomas in other species. The virus has been reported to infect cells from a variety of non-porcine origins and is, therefore, designated as a xenotropic, amphotropic or polytrophic virus (Lieber MM, Sherr CJ. Benveniste RE and Todaro

that it is substantially free of biological material with which the whole genome is normally associated _in vivo_. The isolated polynucleotide fragment may be cloned to provide a recombinant molecule comprising the polynucleotide fragment. Thus, "polynucleotide fragment" includes double and single stranded DNA, RNA and polynucleotide sequences derived therefrom, for example, subsequences of said fragment and which are of any desirable length. Where a nucleic acid is single stranded then both a given strand and a sequence complementary thereto is within the scope of the present invention.

In general, the term "expression product" refers to both transcription and translation products of said polynucleotide fragments. When the expression product is a "polypeptide" (i.e. a chain or sequence of amino acids displaying a biological and/or immunological activity substantially similar to the biological and/or immunological activity of PoEV virion core, polymerase and/or envelope protein), it does not refer to a specific length of the product as such. Thus, the skilled addressee will appreciate that "polypeptide" encompasses _inter alia_ peptides, polypeptides and proteins of PoEV. The polypeptide if required, can be modified _in vivo_ and _in vitro_, for example by glycosylation, amidation, carboxylation, phosphorylation and/or post-translational cleavage.

Polynucleotide fragments comprising portions encompassing the PoEV genome, and derived from retrovirus particles released from a reverse transcriptase-positive porcine kidney cell line PK-15, have been molecularly cloned into a plasmid vector. This was achieved by synthesising cDNAs of PoEV RNA genomes which were

Preferred fragments of this aspect of the invention are polynucleotide fragments encoding: (a) at least one of the three polypeptides having an amino acid sequence which is shown in Figure 1; (b) encoding a polypeptide which is a physiologically active and/or immunogenic derivative of at least one of the polypeptides defined in (a); or (c) which is complementary to a polynucleotide sequence as defined above; or polynucleotide fragments: (a) comprising at least one of the ORFs shown in Figure 1 or comprising a corresponding RNA sequence; (b) comprising a sequence having substantial nucleotide sequence identity with a sequence as described in (a) above; or (c) comprising a sequence which is complementary to a sequence as described in (a) or (b) above. It is to be understood that the term "substantial sequence identity" is taken to mean at least 50% (preferably at least 75%, at least 90%, or at least 95%) sequence identity.

The polynucleotide fragment of the present invention may be used to examine the expression and/or presence of the PoEV virus in donor animals and cells, tissues or organs derived from the donor animals to see if they are suitable for xenotransplantation (i.e. PoEV free). In addition, the recipients of pig cells, tissues or organs can be examined for the presence and/or expression of PoEV virus directly or by co-culture or infection of susceptible detector cells.

A polynucleotide fragment of the present invention may be used to identify polynucleotide sequences within the PoEV genome which are PoEV specific (i.e. it is not necessary for the complete PoEV genome to be identified). Such PoEV specific

there may be less than 10, preferably less than 5 mismatches). Thus, the present invention also provides anti-sense or complementary nucleotide sequence(s) which is/are capable of specifically hybridising to the disclosed DNA sequence. If a PoEV specific polynucleotide is to be used as a primer in PCR and/or sequencing studies, the polynucleotide must be capable of hybridising to PoEV nucleic acid and capable of initiating chain extension from 3' end of the polynucleotide, but not able to correctly initiate chain extension from non PoEV sequences (especially from human, or non-PoEV porcine sequences).

If a PoEV specific test polynucleotide sequence is to be used in hybridisation studies, to test for the presence of PoEV nucleic acid in a sample, the test polynucleotide should preferably remain hybridised to a sample polynucleotide under stringent conditions. If desired, either the test or sample polynucleotide may be immobilised. Generally the test polynucleotide sequence is at least 10 or at least 50 bases in length. It may be labelled by suitable techniques known in the art. Preferably the test polynucleotide sequence is at least 200 bases in length and may even be several kilobases in length. Thus, either a denatured sample or test sequence can be first bound to a support. Hybridization can be effected at a temperature of between 50 and 70°C in double strength SSC (2xNaCl 17.5g/l and sodium citrate (SC) at 8.8g/l) buffered saline containing 0.1% sodium dodecyl sulphate (SDS). This can be followed by rinsing of the support at the same temperature but with a buffer having a reduced SSC concentration. Depending upon the degree of stringency required, and thus the degree of

therein or an RNA equivalent, or with a part of such a strand. Preferably such a part is at least 10, at least 30, at least 50 or at least 200 bases long. It may be an ORF or a part thereof.

Oligonucleotides which are generally greater than 30 bases in length should preferably remain hybridised to a sample polynucleotide under one or more of the stringent conditions mentioned above. Oligonucleotides which are generally less than 30 bases in length should also preferably remain hybridised to a sample polynucleotide but under different conditions of high stringency. Typically the melting temperature of an oligonucleotide less than 30 bases may be calculated according to the formula of; 2°C for every A or T, plus 4°C for every G or C, minus 5°C. Hybridisation may take place at or around the calculated melting temperature for any particular oligonucleotide, in 6 x SSC and 1% SDS. Non specifically hybridised oligonucleotides may then be removed by stringent washing, for example in 3 x SSC and 0.1% SDS at the same temperature. Only substantially similar matched sequences remain hybridised i.e. said oligonucleotide and corresponding PoEV nucleic acid.

When oligonucleotides of generally less than 30 bases in length are used in sequencing and/or PCR studies, the melting temperature may be calculated in the same manner as described above. The oligonucleotide may then be allowed to anneal or hybridise at a temperature around the oligonucleotides calculated melting temperature. In the case of PCR studies the annealing temperature should be around the lower of the calculated melting temperatures for the two priming oligonucleotides. It is to be

(i.e. hybridisation studies) for detecting the presence or otherwise of PoEV polynucleotide in the nucleic acid of pigs or in cell, tissue or organ samples taken from pigs (e.g. from potential transplant organs such as liver, kidney and heart). Such cells, tissues or organs can be derived from transgenic animals produced as described in EP-A-0493852, or by other means known in the art. Thus the cells, tissues or organs of transgenic pigs can be associated with one or more homologous complement restriction factors active in humans to prevent/reduce activation of complement.

Furthermore the polynucleotide fragments of the present invention can be used to analyze the genetic organisation of endogenous PoEV located in the animal cell genome in pigs thus permitting the screening of herds of pigs for altered provirus and genomic loci (e.g. non-expressed provirus loci). Such a screening method would facilitate, for example, screening in a population of animals which are bred to lack expressed provirus and genomic loci and/or loci that do not encode infectious virus particles.

Reagents may also be developed from said polynucleotide fragments as aids to develop pigs that do not express an infectious, PoEV capable of infecting humans. Such pigs could still contain partial defective genomes that could result in the expression of non-infectious particles, viral proteins or viral mRNA. Alternatively, it may be possible to use constructs derived from the PoEV polynucleotide sequence to act as insertional mutagens to knockout the productive infectious PoEV in embryos, embryonic stem cells, or cells containing

taken up by cells and the polynucleotide fragment of interest expressed, producing protein. Presentation of the protein on cell surface stimulates the host immune system to produce antibodies immunoreactive with said protein as part of a defence mechanism. Thus, expressed protein may be used as a vaccine.

Inactivated vaccines can be produced from PoEV's or cells releasing PoEV. Such infected cells may be generated by natural infection or by transfection of a proviral clone of PoEV. It will be understood that a proviral clone is a molecular clone encoding on at least one antigenic polypeptide of PoEV. After harvesting the virus and/or the infected cells, viruses or infected cells present can be inactivated for example, with formaldehyde, gluteraldehyde, acetylethylenimine or other suitable agent or process to generate an inactivated vaccine using methods commonly employed in the art. (CVMP Working Party on Immunological Veterinary Medicinal Products (1993). General requirements for the production and control of inactivated mammalian bacterial and viral vaccines for veterinary use). Sub unit vaccines may be prepared from the individual proteins encoded by the gag, pol and env genes. Typically a vaccine would contain env gene products either alone or in combination with gag genes produced by expression in bacteria, yeast or mammlian cell systems.

Proviral clones of PoEV can be engineered to develop single cycle or replication defective viral vectors suitable for vaccination using techniques. Such viral vectors known in the art (e.g. MuLV Murine Leukaemia Retrovirus, Adenovirus and Herpesviruses (Anderson WF. (1992). Human Gene Therapy. *Science*

the pig family). These variations may be demonstrated by (an) amino acid difference(s) in the overall sequence or by deletions, substitutions, insertions, inversions or additions of (an) amino acid(s) in said sequence. All such derivatives showing active polymerase and/or envelope polypeptide physiological and/or immunological activity are included within the scope of the invention. For example, for the purpose of the present invention conservative replacements may be made between amino acids within the following groups:

(I) Alanine, serine, threonine;

(II) Glutamic acid and aspartic acid;

(III) Arginine and leucine;

(IV) Asparagine and glutamine;

(V) Isoleucine, leucine and valine;

(VI) Phenylalanine, tyrosine and tryptophan

Moreover, recombinant DNA technology may be used to prepare nucleic acid sequences encoding the various derivatives outlined above.

As is well known in the art, the degeneracy of the genetic code permits substitution of bases in a codon resulting in a different codon which is still capable of coding for the same amino acid, e.g. the codon for amino acid glutamic acid is both GAT and GAA. Consequently, it is clear that for the expression of polypeptides with the amino acid sequences shown in Figure 1 or fragments thereof, use can be made of a derivative nucleic acid sequence with such an alternative codon composition different from the nucleic acid sequence shown in said Figure 1.

Furthermore, fragments derived from the PoEV polymerase

sequences according to the invention are known in the art (e.g. Rodriguez, R.L. and Denhadt, D.T., Edit., Vectors: a survey of molècular cloning vectors and their uses, Butterworths, 1988).

The methods to be used for the construction of a recombinant nucleic acid molecule according to the invention are known to those of ordinary skill in the art and are inter alia set forth in Sambrook, et al. (Molecular Cloning: a laboratory manual Cold Spring Harbour Laboratory, 1989).

The present invention also relates to a transformed cell containing the PoEV polynucleotide fragment in an expressible form. "Transformation", as used herein, refers to the introduction of a heterologous polynucleotide fragment into a host cell. The method used may be any known in the art, for example, direct uptake, transfection or transduction (Current Protocols in Molecular Biology, 1995. John Wiley and Sons Inc). The heterologous polynucleotide fragment may be maintained through autonomous replication or alternatively, may be integrated into the host genome. The recombinant nucleic acid molecules preferably are provided with appropriate control sequences compatible with the designated host which can regulate the expression of the inserted polynucleotide fragment, e.g. tetracycline responsive promoter, thymidine kinase promoter, SV-40 promoter and the like.

Suitable hosts for the expression of recombinant nucleic acid molecules may be prokaryotic or eukaryotic in origin. Hosts suitable for the expression of recombinant nucleic acid molecules may be selected from bacteria, yeast, insect cells and mammalian cells.

NaCl,1mM EDTA) and the virions disrupted by the addition of 1ml of lysis/binding buffer. Dynabeads Oligo$(dT)_{25}$ were conditioned according to the manufacturer's instructions and added to the virus disrupted solution. Viral RNA is allowed to bind to the Dynabead for 10 minutes before the supernatant is removed and the bound RNA was washed three times with washing buffer with LiDS (0.5ml) and twice with washing buffer alone. The RNA is finally resuspended in 25 μl of elution solution. All procedures were performed at ambient temperature. RNase contamination was avoided by the wearing of gloves, observation of sterile technique and treatment of solutions and non-disposable glass and plasticware with diethyl pyrocarbonate (DEPC). The RNA was resuspended in DEPC- treated sterile water.

Example 2

Synthesis of cDNA

cDNA was synthesised from the purified genomic RNA using Great Lengths ™ cDNA amplification reverse transcriptase reagents (Clonetech Laboratories Inc.) following the manufacturer's instructions. The RNA was primed with both oligo(dT) and random hexamers to maximise synthesis.

The Great Lengths cDNA synthesis protocol is based on a modified Gubler and Hoffman (1983) protocol for generating complementary DNA libraries and essentially consists of first-strand synthesis, second strand synthesis, adaptor ligation, and size fractionaction.

First strand synthesis: lock-docking primers anneal to the

Example 3

Molecular cloning of cDNA

The size fractionated fragment was ligated with EcoR I- digested pZErO™ -1 plasmid vector DNA (Invitrogen Corporation, San Diego, U.S.). The ligation mix was used to transform competent TOP10F'cells and these were plated onto L-Agar containing zeocin following the manufacturer's instructions (Zero Background™ cloning kit - Invitrogen). Several of the resulting zeocin resistant colonies were amplified in L-Broth containing zeocin and the plasmid DNA was purified by alkaline lysis (Maniatis *et al.*, 1982).

The plasmid DNA was digested to completion with the endonuclease EcoR I and the resulting DNA fragments were separated by electrophoresis through an 1.0% agarose gel (Maniatis et *al.*, 1982), in order to check that a fragment in the predicted size fractionated size range had been cloned. A clone identified as pPoEV was used in further experimentation.

Example 4

DNA sequence analysis.

pPoEV plasmid DNA was purified according to common techniques (Sambrook et al, 1989) and sequenced using an ABI automated sequencer. Overlapping sequencing primers from both strands of the molecular clone were used to determine the nucleotide sequence. Homologies were observed between pPoEV and the majority of retroviruses determined by using alogrithims from DNASTAR Inc. Lasergene software (DNASTAR). The homologies were closest with

15 minutes. The process was repeated one further time. The sample was mixed with 5ml (3x volume) of extraction buffer (Maniatis *et al.*, 1982).

Purification

The samples (i.e. cultured cells, porcine tissue or porcine blood cells) in proteinase K-extraction buffer containing 20μg/ml RNAse and 100μg/ml proteinase K were digested for approximately 24 hours at 37°C. The deproteinised DNA was extracted twice with phenol and twice with phenol chloroform and finally precipitated by ethanol in the presence of ammonium acetate. The DNA was recovered by centrifugation at 3000g for 30 minutes and the supernatant discarded (Maniatis *et al.*, 1982). The pellet was washed in 70% ethanol and allowed to air dry for approximately 1 hour. The DNA was allowed to re-dissolve in Tris EDTA (TE) buffer and the purity and concentration of the DNA was assessed by spectrophotometry (Maniatis *et al.*, 1982).

Example 6

Southern blot analysis of porcine tissue and cells

In order to demonstrate that the molecularly cloned DNA comprising the insert from PoEV was derived from the PK-15 cell line (American Type Culture Collection CCL33), the DNA was hybridised against cellular DNAs and its ability to detect proviral DNA was examined.

DNA purified from pPoEV was radioactively labelled and used to probe a Southern blot of endonuclease digested DNAs derived from PK-15 cells .

Culture Collection CRL 1746) and primary porcine kidney cells (Central Veterinary Laboratory batch C04495) but not in hamster CHOK1 (American Type Culture Collection CCL61) or murine NSO myeloma cells (European Collection of Animal Cell Cultures 85110503).

In order to demonstrate that the molecularly cloned DNA comprising the insert from pPoEV could detect sequences in porcine cells and tissues in addition to PK-15 the pPoEV DNA was hybridised against cellular DNA from tissues derived from pigs and its ability to detect proviral DNA was examined (Maniatis *et al.*, 1982).

The DNA purified from pPoEV was radioactively labelled and used to probe a Southern blot of endonuclease digested DNAs derived from pig organs including liver, kidney, heart and blood.

The DNAs probed were as follows :

a) Copy number controls of pPoEV DNA linearized by digestion with EcoRI. One copy per haploid cell genome was estimated to be 6.84pg. The control was present at an estimated copy number of 5,10, 20 and 50 copies.

b) DNA purified from the porcine tissues digested with EcoRI.

A hybridisation signal was observed in all the porcine DNAs.

On completion of the amplification, 10μl of the reaction mixture was rophoresed through a 5 per cent acrylamide gel. The DNA was visualised by staining with ethidium bromide and exposure to ultraviolet light (320nm).

The PCR reaction amplified a sequence of approximately 787bp from pPoEV and from porcine cells as expected indicating that the assay detected the PoEV proviral DNA. There was no specific amplification of the expected sequence in cells of non-porcine origin and therefore, the PCR reaction and recombinant clone can be used as a specific and sensitive diagnostic tool for detection of PoEV.

Example 8

Production of PoEV polypeptide in *Escherichia coli*.

The open reading frame (ORF) encoding the *pol* peptide was isolated from the pPoEV clone and molecularly cloned into the plasmid pGEX-4T-1 (Pharmacia Ltd.) for expression.

Two ml cultures of *E. coli* transformed with various expression constructs were grown with shaking at 37°C to late log phase ($O.D._{600nm}$ of 0.6) and induced by the addition of IPTG to 0.1 mM. Induced cultures were then incubated for a further 2 hours after which the bacteria were collected by centrifugation. The bacterial pellet was lysed by boiling in SDS-PAGE sample buffer and the protein profile of the induced bacteria was analysed on a 12% acrylamide gel (Laemmli, 1970) followed by staining with coomassie brilliant blue dye.

(a) at least one of the three polypeptides having an amino acid sequence which is shown in Figure 1;

(b) encoding a polypeptide which is a physiologically active and/or immunogenic derivative of at least one of the polypeptides defined in (a); or

(c) which is complementary to a polynucleotide sequence as defined in (a) or (b).

5. As isolated polynucleotide fragment according to any one of claims 1 to 4;

(a) comprising at least one of the ORFs shown in Figure 1 or comprising a corresponding RNA sequence;

(b) comprising a sequence having substantial nucleotide sequence identity with a sequence as described in (a) above; or

(c) comprising a sequence which is complementary to a sequence as described in (a) or (b) above.

6. A recombinant nucleic acid molecule comprising a polynucleotide fragment according to any one of claims 1 to 5.

7. A recombinant nucleic acid molecule according to claim 6 wherein the recombinant nucleic acid molecule comprises regulatory control sequences operably linked to said polynucleotide fragment for controlling expression of said polynucleotide fragment.

17. A probe or a primer according to claim 15 or 16 which has substantial nucleotide sequence identity with a strand of the molecule depicted in Figure 1, or a strand complementary therewith, with a corresponding RNA molecule, or with a part of such a molecule.

18. A PoEV detection kit comprising a polynucleotide primer or probe according to any of claims 15 to 17.

19. Use of a PoEV specific polynucleotide in the detection of PoEV in a sample.

20. Use of a PoEV specific polynucleotide in a PCR for the detection of PoEV in a sample.

21. A pig modified so as to not express an infectious PoEV capable of infecting humans.

22. Cells, tissues or organs obtainable from a pig encoding to claim 21.

23. Use of a recombinant PoEV polypeptide according to either of claims 11 or 12 in the preparation of a vaccine.

24. Use of a polynucleotide primer or probe according to any of claims 15 to 17 in the preparation of a detection kit capable of detecting the presence of PoEV nucleic acid in a sample.

ABSTRACT

PORCINE RETROVIRUS

The present invention provides porcine retrovirus (PoEV) polynucleotide fragments, particularly those encoding at least one PoEV expression product, a recombinant vector comprising such a polynucleotide fragment or fragments, use of PoEV polynucleotide fragments in the detection of native PoEV, a host cell containing at least one PoEV polynucleotide fragment or recombinant vector, PoEV polypeptides, antibodies immuno-reactive with PoEV polypeptides, pharmaceutical compositions comprising recombinant PoEV polypeptides for use as prophylactic and/or therapeutic agents and uses of PoEV polynucleotide fragments and/or polypeptides in medicine, including veterinary medicine and in the preparation of medicaments for use in medicine.

Figure 1

```
  1  TGTGGGCCCCAGCGCGCTTGG  GAAAAATCCTCTTGCTGTTTGCATCAAGACCGCTTCT   60

 61  CGTGAGTGATTTGGGGTGTCG  GTCTTCCGAGCCCGGACCGACGGGGATTGCTCTTTTACT  120

121  GGCCTTTCATTTGGTGCGTTC  CGGGAAAATCCTGCGACCACCCCTTACAGCCGAGAACC  180

181  GACTTGGAGGTAAAGGGATCC  TTTGGAACGTGTGTGTGTGTCGGCCGGCGTCCCTGTT  240

241  CTGAGTGTCTGTTTTCGGTGA GCGCGCTTTCGGTTTGCAGCTGTCCTCTCAGACCGTAA  300

301  GGACTGGAGGACTGTGATCAG  GACGTGCTAGGAGGATCACAGGCTGCCACCCTGGGG  360

361  ACGCCCCGGGAGGTGGGGAGA  CAGGGACGCCTCGTGGTCTCCTACTGTCGGTCAGAGG  420

421  ACCGAGTTCTGTTGTTGAAGC  AAGCTTCCCCCTCCGCGGCCGTCCGACTCTTTTGCCT  480

481  GCTTGTGGAAGACGCGGACGG  CGCGTGTGTCTGGATCTGTTGGTTTCTCTCTCGTGTG  540

541  TCTTTGTCTTGTGCGTCCTTG  TACAGTTTTAATATGGGACAGACAGTGACTACCCCCC  600
                                             MetGlyGlnThrValThrThrProL

601  TTAGTTTGACTCTCGACCATT  ACTGAAGTTAGATCCAGGGCTCATAATTTGTCAGTTC  660
     euSerLeuThrLeuAspHisT  hrGluValArgSerArgAlaHisAsnLeuSerValG

661  AGGTTAAGAAGGGACCTTGGC  ACTTTCTGTGCCTCTGAATGGCCAACATTCGATGTTG  720
     lnValLysLysGlyProTrpG  ThrPheCysAlaSerGluTrpProThrPheAspValG

721  GATGGCCATCAGAGGGCACCT  AATTCTGAAATTATCCTGGCTGTTAAGGCAATCATTT  780
     lyTrpProSerGluGlyThrP  AsnSerGluIleIleLeuAlaValLysAlaIleIleP

781  TTCAGACTGGACCCGGCTCTC  CCTGATCAGGAGCCCTATATCCTTACGTGGCAAGATT  840
     heGlnThrGlyProGlySerP  ProAspGlnGluProTyrIleLeuThrTrpGlnAspL

841  TGGCAGAAGATCCTCCGCCAT  GTTAAACCATGGCTAAATAAACCAAGAAAGCCAGGTC  900
     euAlaGluAspProProProT  ValLysProTrpLeuAsnLysProArgLysProGlyP

901  CCCGAATCCTGGCTCTTGGAG  AAAAACAAACACTCGGCCGAAAAAGTCGAGCCCTCTT  960
     roArgIleLeuAlaLeuGlyG  LysAsnLysHisSerAlaGluLysValGluProSerS

961  CCTCGTATCTACCCCGAGATC  GGAGCCGCCGACTTGGCCGGAACCCCACCCTGTTCCC 1020
```

2/9

```
      erSerTyrLeuProArgAsnP GlyAlaAlaAspLeuAlaGlyThrProThrCysSerP

1021  CCACCCCCTTATCCAGCACAC TGCTCTGAGGGGACCTCTGCCCCTCCTGGAGCTCCGG    1080
      roThrProLeuSerSerThrC ysCysGluGlyThrSerAlaProProGlyAlaProV

1081  TGGTGGAGGGACCTGCTGCCG ACTCGGAGCCGGAGAGGCGGCCACCCCGGAGCGGACAG   1140
      alValGluGlyProAlaAlaG ThrArgSerArgArgGlyAlaThrProGluArgThrA

1141  ACGAGATCGCGATATTACCGC GCACCTATGGCCCTCCCATGCCAGGGGGCCAATTGC    1200
      spGluIleAlaIleLeuProA rgThrTyrGlyProProMetProGlyGlyGluLeuG

1201  AGCCCCTCCAGTATTGGCCCA TCTTCTGCAGATCTCTATAATTGGAAGACTAACCATC   1260
      lnProLeuGlnTyrTrpProP roSerSerAlaAspLeuTyrAsnTrpLysThrAsnHisP

1261  CCCCTTTCTCGGAGGATCCCC GCCTCACGGGGTTGGTGGAGTCCCTTATGTTCTCTC    1320
      roProPheSerGluAspProG ArgLeuThrGlyLeuValGluSerLeuMetPheSerH

1321  ACCAGCCTACTTGGGATGATT CAACAGCTGCTGCAGACACTCTTCACAACCGAGGAGC   1380
      isGlnProThrTrpAspAspC GlnGlnLeuLeuGlnThrLeuPheThrThrGluGluA

1381  GAGAGAGAATTCTGTTAGAGG AGAAAAATGTTCCTGGGGCCGACGGGCGACCCACGC    1440
      rgGluArgIleLeuLeuGluA ArgLysAsnValProGlyAlaAspGlyArgProThrC

1441  AGTTGCAAAATGAGATTGACA GGATTTCCCTTGACTCGCCCCGGTTGGGACTACAACA   1500
      laLeuGlnAsnGluIleAspA GlyPheProLeuThrArgProGlyTrpAspTyrAsnT

1501  CGGCTGAAGGTAGGGAGAGCT AAAATCTATCGCCAGGCTCTGGTCGCCGGTCTCCGGG   1560
      hrAlaGluGlyArgGluSerL LysIleTyrArgGlnAlaLeuValAlaGlyLeuArgG

1561  GCGCCTCAAGACGGCCCACTA TTGGCTAAGGTAAGAGAGGTGATGCAGGGACCGAACG   1620
      lyAlaSerArgArgProThrA LeuAlaLysValArgGluValMetGlnGlyProAsnG

1621  AACCTCCCTCGGTATTTCTTC AGGCTCATGGAAGCCTTCAGGCGGTTCACCCCTTTTG   1680
      luProProSerValPheLeuG ArgLeuMetGluAlaPheArgArgPheThrProPheA

1681  ATCCTACCTCAGAGGCCCAGA GCCTCAGTGGCCCTGGCCTTCATTGGGCAGTCGGCTC   1740
      spProThrSerGluAlaGlnL AlaSerValAlaLeuAlaPheIleGlyGlnSerAlaL

1741  TGGATATCAGGAAGAAACTTC AGACTGGAAGGGTTACAGGAGGCTGAGCTACGTGATC   1800
      euAspIleArgLysLysLeuG ArgLeuGluGlyLeuGlnGluAlaGluLeuArgAspL

1801  TAGTGAGAGAGGCAGAGAAGG TATTACAGAAGGGAGACAGAAGAGGAGAAGGAACAGA   1860
      euValArgGluAlaGluLysV TyrTyrArgArgGluThrGluGluGluLysGluGlnA

1861  GAAAAGAAAAGGAGACAGAAC GGGAGGAAAGACGTGATAGACGGCAAGAGAAGAATT    1920
      rgLysGluLysGluArgGluA ArgGluGluArgArgAspArgArgGlnGluLysAsnL
```

3/9

```
1921  TGACTAAGATCTTGGCCGCAG GTTGAAGGGAACAGCAGCAGGGAGAGAAGAGAGATT  1980
      euThrLysIleLeuAlaAlaV ValGlnGlyLysSerSerArgGluArgGluArgAspP

1981  TTAGGAAAATTAGGTCAGGCA AGACAGTCAGGGAACCTGGGCAATAGGACCCCACTCG  2040
      heArgLysIleArgSerGlyA ArgGlnSerGlyAsnLeuGlyAsnArgThrProLeuA

2041  ACAAGGACCAGTGTGCGTATT AAAGAAAAAGGACACTGGGCAAGGAACTGCCCCAAGA  2100
      spLysAspGlnCysAlaTyrC LysGluLysGlyHisTrpAlaArgAsnCysProLysL

2101  AGGGAAACAAAGGACCGAAGT TAGCTCTAGAAGAAGATAAAGATTAGGTGAGACGGGT  2160
      ysGlyAsnLysGlyProLysP End                   ArgLeuGlyGluThrGly

2161  TCGGACCCCCTCCCCGAGCCC GGTAACTTTGAAGGTGGAGGGGCAACCAGTTGAGTTC  2220
      SerAspProLeuProGluPro rgValThrLeuLysValGluGlyGlnProValGluPhe

2221  CTGGTTGATACCGGAGCGGAG TTCAGTCCTGCTACAACCATTAGGAAAACTAAAAGAA  2280
      LeuValAspThrGlyAlaGlu sSerValLeuLeuGlnProLeuGlyLysLeuLysGlu

2281  AAAAAATCCTGGGTGATGGGT CACAGGGCAACGGCAGTATCCATGGACTACCCGAAGA  2340
      LysLysSerTrpValMetGly aThrGlyGlnArgGlnTyrProTrpThrThrArgArg

2341  ACCGTTGACTTGGGAGTGGGA GGTAACCCACTCGTTTCTGGTCATCCCTGAGTGCCCA  2400
      ThrValAspLeuGlyValGly gValThrHisSerPheLeuValIleProGluCysPro

2401  GTACCCCTTCTAGGTAGAGAC ACTCACCAAGATGGGAGCTCAAATTTCTTTTGAACAA  2460
      ValProLeuLeuGlyArgAsp uLeuThrLysMetGlyAlaGlnIleSerPheGluGln

2461  GGAAGACCAGAAGTGTCTGTG AACAAACCCATCACTGTGTTGACCCTCCAATTAGAT  2520
      GlyArgProGluValSerVal nAsnLysProIleThrValLeuThrLeuGlnLeuAsp

2521  GATGAATATCGACTATATTCT CAAGTAAAGCCTGATCAAGATATACAGTCCTGGTTG  2580
      AspGluTyrArgLeuTyrSer rGlnValLysProAspGlnAspIleGlnSerTrpLeu

2581  GAGCAGTTTCCCCAAGCCTGG AGAAACGCAGGGATGGGTTTGGCAAAGCAAGTTCCC  2640
      GluGlnPheProGlnAlaTrp uGluThrAlaGlyMetGlyLeuAlaLysGlnValPro

2641  CCACAGGTTATTCAACTGAAG AGTGCAACACCAGTATCAGTCAGACAGTACCCCTTG  2700
      ProGlnValIleGlnLeuLys sSerAlaThrProValSerValArgGlnTyrProLeu

2701  AGTAGAGAGGCTCGAGAAGGA TTGGCCGCATGTTCAAAGATTAATCCAACAGGGCATC  2760
      SerArgGluAlaArgGluGly pTrpProHisValGlnArgLeuIleGlnGlnGlyIle

2761  CTAGTTCCTGTCCAATCCCCT AATACTCCCCTGCTACCGGTTAGGAAACCTGGGACC  2820
      LeuValProValGlnSerPro nAsnThrProLeuLeuProValArgLysProGlyThr

2821  AATGATTATCGACCAGTACAT CTTGACAGAGGTCAATAAAAGGGTGCACGACATACAC  2880
```

AsnAspTyrArgProValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHis

2881 CCAACGGTCCCGAACCCTTATAACCTCTTGAGCGCCCTCCCGCCTGAACGGAACTGGTAC 2940
ProThrValProAsnProTyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyr

2941 ACAGTATTGGACTTAAAAGATGCCTTCTTCTGCCTGAGATTACACCCCACTAGCCAACCA 3000
ThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnPro

3001 CTTTTTGCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAGCTCACCTGGACC 3060
LeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThr

3061 CGACTGCCCCAAGGGTTCAAGAACTCCCCGACCATCTTTGACGAAGCCCTACACAGGGAC 3120
ArgLeuProGlnGlyPheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAsp

3121 CTGGCCAACTTCAGGATCCAACACCCTCAGGTGACCCTCCTCCAGTACGTGGATGACCTG 3180
LeuAlaAsnPheArgIleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeu

3181 CTTCTGGCGGGAGCCACCAAACAGGACTGCTTAGAAGGTACGAAGGCACTACTGCTGGAA 3240
LeuLeuAlaGlyAlaThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGlu

3241 TTGTCTGACCTAGGCTACAGAGCCTCTGCTAAGAAGGCCCAGATTTGCAGGAGAGAGGTA 3300
LeuSerAspLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluVal

3301 ACATACTTGGGGTACAGTTTGCGGGGCGGGCAGCGATGGCTGACGGAGGCACGGAAGAAA 3360
ThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGluAlaArgLysLys

3361 ACTGTAGTCCAGATACCGGCCCCAACCACAGCCAAACAAGTGAGAGAGTTTTTGGGGACA 3420
ThrValValGlnIleProAlaProThrThrAlaLysGlnValArgGluPheLeuGlyThr

3421 GCTGGATTTTGCAGACTGTGGATCCCGGGGTTTGCGACCTTAGCAGCCCCACTCTACCCG 3480
AlaGlyPheCysArgLeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyrPro

3481 CTAACCAAAGAAAAAGGGGGATTCTCCTGGGCTCCTGAGCACCAGAAGGCATTTGATGCT 3540
LeuThrLysGluLysGlyGlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAla

3541 ATCAAAAAGGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACTAAACCCTTT 3600
IleLysLysAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPhe

3601 ACCCTTTATGTGGATGAGCGTAAGGGAGTAGCCCGAGGAGTTTTAACCCAAACCCTAGGA 3660
ThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGly

3661 CCATGGAGGAGACCTGTTGCCTACCTGTCAAAGAAGCTTGATCCTGTAGCCAGTGGTTGG 3720
ProTrpArgArgProValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrp

3721 CCCGTATGTCTGAAGGCTATCGCAGCTGTGGCCATACTGGTCAAGGACGCTGACAAATTG 3780
ProValCysLeuLysAlaIleAlaAlaValAlaIleLeuValLysAspAlaAspLysLeu

5/9

3781 ACTTTGGGACAGAATATAACTGTAATAGCCCCCCATGCATTGGAGAACATCGTTCGGCAG 3840
ThrLeuGlyGlnAsnIleThrValIleAlaProHisAlaLeuGluAsnIleValArgGln

3841 CCCCCAGACCGATGGATGACCAACGCCCGCATGACCCACTATCAAAGCCTGCTTCTCACA 3900
ProProAspArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThr

3901 GAGAGGGTCACTTTCGCTCCACCAGCCGCTCTCAACCCTGCCACTCTTCTGCCTGAAGAG 3960
GluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluGlu

3961 ACTGATGAACCAGTGACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGGGTCCGC 4020
ThrAspGluProValThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyValArg

4021 AAGGACCTTACAGACATACCGCTGACTGGAGAAGTGCTAACCTGGTTCACTGACGGAAGC 4080
LysAspLeuThrAspIleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySer

4081 AGCTATGTGGTGGAAGGTAAGAGGATGGCTGGGGCGGCAGTGGTGGACGGGACCCGCACG 4140
SerTyrValValGluGlyLysArgMetAlaGlyAlaAlaValValAspGlyThrArgThr

4141 ATCTGGGCCAGCAGCCTGCCGGAAGGAACTTCAGCGCAAAAGGCTGAGCTCATGGCCCTC 4200
IleTrpAlaSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeu

4201 ACGCAAGCTTTGCGGCTGGCCGAAGGGAAATCCATAAACATTTATACGGACAGCAGGTAT 4260
ThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyr

4261 GCCTTTGCGACTGCACACGTACACGGGGCCATCTATAAACAAAGGGGGTTGCTTACCTCA 4320
AlaPheAlaThrAlaHisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSer

4321 GCAGGGAGGGAAATAAAGAACAAAGAGGAAATTCTAAGCCTATTAGAAGCCTTACATTTG 4380
AlaGlyArgGluIleLysAsnLysGluGluIleLeuSerLeuLeuGluAlaLeuHisLeu

4381 CCAAAAAGGCTAGCTATTATACACTGTCCTGGACATCAGAAAGCCAAAGATCTCATATCT 4440
ProLysArgLeuAlaIleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSer

4441 AGAGGGAACCAGATGGCTGACCGGGTTGCCAAGCAGGCAGCCCAGGCTGTTAACCTTCTG 4500
ArgGlyAsnGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuLeu

4501 CCTATAATAGAAACGCCCAAAGCCCCAGAACCCAGACGACAGTACACCCTAGAAGACTGG 4560
ProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThrLeuGluAspTrp

4561 CAAGAGATAAAAAAGATAGACCAGTTCTCTGAGACTCCGGAGGGGACCTGCTATACCTCA 4620
GlnGluIleLysLysIleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSer

4621 TATGGGAAGGAAATCCTGCCCCACAAAGAAGGGTTAGAATATGTCCAACAGATACATCGT 4680
TyrGlyLysGluIleLeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArg

4681 CTAACCCACCTAGGAACTAAACACCTGCAGCAGTTGGTCAGAACATCCCCTTATCATGTT 4740

```
      LeuThrHisLeuGlyThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisVal

4741  CTGAGGCTACCAGGAGTGGCTGACTCGGTGGTCAAACATTGTGTGCCCTGCCAGCTGGTT  4800
      LeuArgLeuProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuVal

4801  AATGCTAATCCTTCCAGAATACCTCCAGGAAAGAGACTAAGGGGAAGCCACCCAGGCGCT  4860
      AsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySerHisProGlyAla

4861  CACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAAATACGGAAACAAATATCTATTG  4920
      HisTrpGluValAspPheThrGluValLysProAlaLysTyrGlyAsnLysTyrLeuLeu

4921  GTTTTTGTAGACACCTTTTCAGGATGGGTAGAGGCTTATCCTACTAAGAAAGAGACTTCA  4980
      ValPheValAspThrPheSerGlyTrpValGluAlaTyrProThrLysLysGluThrSer

4981  ACCGTGGTGGCTAAGAAAATACTGGAGGAAATTTTTCCAAGATTTGGAATACCTAAGGTA  5040
      ThrValValAlaLysLysIleLeuGluGluIlePheProArgPheGlyIleProLysVal

5041  ATAGGGTCAGACAATGGTCCAGCTTTCGTTGCCCAGGTAAGTCAGGGACTGGCCAAGATA  5100
      IleGlySerAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIle

5101  TTGGGGATTGATTGGAAACTGCATTGTGCATACAGACCCCAAAGCTCAGGACAGGTAGAG  5160
      LeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGlu

5161  AGGATGAATAGAACCATTAAAGAGACCCTTACCAAATTGACCACAGAGACTGGCATTAAT  5220
      ArgMetAsnArgThrIleLysGluThrLeuThrLysLeuThrThrGluThrGlyIleAsn

5221  GATTGGATGGCTCTCCTGCCCTTTGTGCTTTTTAGGGTGAGGAACACCCCTGGACAGTTT  5280
      AspTrpMetAlaLeuLeuProPheValLeuPheArgValArgAsnThrProGlyGlnPhe

5281  GGGCTGACCCCCTATGAATTGCTCTACGGGGGACCCCCCCCGTTGGCAGAAATTGCCTTT  5340
      GlyLeuThrProTyrGluLeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPhe

5341  GCACATAGTGCTGATGTGCTGCTTTCCCAGCCTTTGTTCTCTAGGCTCAAGGCGCTCGAG  5400
      AlaHisSerAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGlu

5401  TGGGTGAGGCAGCGAGCGTGGAAGCAGCTCCGGGAGCCCTACTCAGGAGGAGACTTGCAA  5460
      TrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGlyGlyAspLeuGln

5461  GTTCCACATCGCTTCCAAGTTGGAGATTCAGTCTATGTTAGACGCCACCGTGCAGGAAAC  5520
      ValProHisArgPheGlnValGlyAspSerValTyrValArgArgHisArgAlaGlyAsn

5521  CTCGAGACTCGGTGGAAGGGACCTTATCTCGTACTTTTGACCACACCAACGGCTGTGAAA  5580
      LeuGluThrArgTrpLysGlyProTyrLeuValLeuLeuThrThrProThrAlaValLys

5581  GTCGAAGGAATCCCCACCTGGATCCATGCATCCCACGTTAAGCCGGCGCCACCTCCCGAC  5640
      ValGluGlyIleProThrTrpIleHisAlaSerHisValLysProAlaProProProAsp
                                MetHisProThrLeuSerArgArgHisLeuProTh
```

7/9

```
5641  TCGGGGTGGAGAGCCGAAAAGACTGAGAATCCCCTTAAGCTTCGCCTCCATCGCCTGGTT  5700
      SerGlyTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgLeuVal
      rArgGlyGlyGluProLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPh

5701  CCTTACTCTAACAATAACTCCCCAGGCCAGTAGTAAACGCCTTATAGACAGCTCGAACCC  5760
      ProTyrSerAsnAsnAsnSerProGlyGlnEnd
      eLeuThrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSerAsnPr

5761  CCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCCTGATACGGGTGTCACTGTAAA  5820
      oHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspThrGlyValThrValAs

5821  TAGCACTCGAGGTGTTGCTCCTAGAGGCACCTGGTGGCCTGAACTGCATTTCTGCCTCCG  5880
      nSerThrArgGlyValAlaProArgGlyThrTrpTrpProGluLeuHisPheCysLeuAr

5881  ATTGATTAACCCCGCTGTTAARAGCACACCTCCCAACCTAGTCCGTAGTTATGGGTTCTA  5940
      gLeuIleAsnProAlaValLysSerThrProProAsnLeuValArgSerTyrGlyPheTy

5941  TTGCTGCCCAGGCACAGAGAAAGAGAAATACTGTGGGGGTTCTGGGGAATCCTTCTGTAG  6000
      rCysCysProGlyThrGluLysGluLysTyrCysGlyGlySerGlyGluSerPheCysAr

6001  GAGATGGAGCTGCGTCACCTCCAACGATGGAGACTGGAAATGGCCGATCTCTCTCCAGGA  6060
      gArgTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeuGlnAs

6061  CCGGGTAAAATTCTCCTTTGTCAATTCCGGCCCGGGCAAGTACAAAATGATGAAACTATA  6120
      pArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLysMetMetLysLeuTy

6121  TAAAGATAAGAGCTGCTCCCCATCAGACTTAGATTATCTAAAGATAAGTTTCACTGAAAG  6180
      rLysAspLysSerCysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluAr

6181  GAAAACAGGAAAATATTCAAAAGTGGATAAATGGTATGAGCTGGGGAATAGTTTTTTATT  6240
      gLysThrGlyLysTyrSerLysValAspLysTrpTyrGluLeuGlyAsnSerPheLeuLe

6241  ATATGGCGGGGGAGCAGGGTCCACTTTAACCATTCGCCTTAGGATAGAGACGGGGACAGA  6300
      uTyrGlyGlyGlyAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGl

6301  ACCCCCTGTGGCAATGGGACCCGATAAAGTACTGGCTGAACAGGGGCCCCCGGCCCTGGA  6360
      uProProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProProAlaLeuGl

6361  GCCACCGCATAACTTGCCGGTGCCCCAATTAACCTCGCTGCGGCCTGACATAACACAGCC  6420
      uProProHisAsnLeuProValProGlnLeuThrSerLeuArgProAspIleThrGlnPr

6421  GCCTAGCAACAGTACCACTGGATTGATTCCTACCAACACGCCTAGAAACTCCCCAGGTGT  6480
      oProSerAsnSerThrThrGlyLeuIleProThrAsnThrProArgAsnSerProGlyVa

6481  TCCTGTTAAGACAGGACAGAGACTCTTCAGTCTCATCCAGGGAGCTTTCCAAGCCATCAA  6540
      lProValLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAs
```

```
6541 CTCCACCGACCCTGATGCCACTTCTTCTTGTTGGCTTTGTCTATCCTCAGGGCCTCCTTA 6600
     nSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTy

6601 TTATGAGGGGATGGCTAAAGAAAGAAAATTCAATGTGACCAAAGAGCATAGAAATCAATG 6660
     rTyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArgAsnGlnCy

6661 TACATGGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTTCCGGGAAGGGGACATGCAT 6720
     sThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIl

6721 AGGAAAAGCTCCCCCATCCCACCAACACCTTTGCTATAGTACTGTGGTTTATGAGCAGGC 6780
     eGlyLysAlaProProSerHisGlnHisLeuCysTyrSerThrValValTyrGluGlnAl

6781 CTCAGAAAATCAGTATTTAGTACCTGGTTATAACAGGTGGTGGCATGCAATACTGGGTT 6840
     aSerGluAsnGlnTyrLeuValProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLe

6841 AACCCCCTGTGTTTCCACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTCATGGTCCA 6900
     uThrProCysValSerThrSerValPheAsnGlnSerLysAspPheCysValMetValGl

6901 AATCGTCCCCCGAGTGTACTACCATCCTGAGGAAGTGGTCCTTGATGAATATGACTATCG 6960
     nIleValProArgValTyrTyrHisProGluGluValValLeuAspGluTyrAspTyrAr

6961 GTATAACCGACCAAAAAGAGAACCCGTATCCCTTACCCTAGCTGTAATGCTCGGATTAGG 7020
     gTyrAsnArgProLysArgGluProValSerLeuThrLeuAlaValMetLeuGlyLeuGl

7021 GACGGCCGTTGGCGTAGGAACAGGGACAGCTGCCCTGATCACAGGACCACAGCAGCTAGA 7080
     yThrAlaValGlyValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGl

7081 GAAAGGACTTGGTGAGCTACATGCGGCCATGACAGAAGATCTCCGAGCCTTAAAGGAGTC 7140
     uLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysGluSe

7141 TGTTAGCAACCTAGAAGAGTCCCTGACTTCTTTGTCTGAAGTGGTTCTACAGAACCGGAG 7200
     rValSerAsnLeuGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgAr

7201 GGGATTAGATCTGCTGTTTCTAAGACAAGGTGGGTTATGTGCAGCCTTAAAAGAAGAATG 7260
     gGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLysGluGluCy

7261 TTGCTTCTATGTAGATCACTCAGGAGCCATCAGAGACTCCATGAACAAGCTTAGAAAAAA 7320
     sCysPheTyrValAspHisSerGlyAlaIleArgAspSerMetAsnLysLeuArgLysLy

7321 GTTAGAGAGGCGTCGAAGGGAAAGAGAGGCTGACCAGGGGTGGTTTGAAGGATGGTTCAA 7380
     sLeuGluArgArgArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAs

7381 CAGGTCTCCTTGGATGACCACCCTGCTTTCTGCTCTGACGGGGCCCCTAGTAGTCCTGCT 7440
     nArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeuLe
```

7441 CCTGTTACTTACAGTTGGGCCTTGCTTAATTAATAGGTTTGTTGCCTTTGTTAGAGAACG 7500
uLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluAr

7501 AGTGAGTGCAGTCCAGATCATGGTACTTAGGCAACAGTACCAAGGCCTTCTGAGCCAAGG 7560
gValSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeuSerGlnGl

7561 AGAAACTGACCTCTAGCCTTCCCAGTTCTAAGATTAGAACTATTAACAAGACAAGAAGTG 7620
yGluThrAspLeuEnd

7621 GGGAATGAAAGGATGAAAATGCAACCTAACCCTCCCAGAACCCAGGAAGTTAATAAAAAG 7680

7681 CTCTAAATGCCCCCGAATTMCAGACCCTGCTGGCTGCCAGTAAATAGGTAGAAGGTCACA 7740

7741 CTTCCTATTGTTCCAGGGCCTGCTATCCTGGCCTAAGTAAGATAACAGGAAATGAGTTGA 7800

7801 CTAATCGCTTATCTGGATTCTGTAAAACTGACTGGCACCATAGAAGAATTGATTACACAT 7860

7861 TGACAGCCCTAGTGACCTATCTCAACTGCAATCTGTCACTCTGCCCAGGAGCCCACGCAG 7920

7921 ATGCGGACCTCCGGAGCTATTTTAAAATGATTGGTCCACGGAGCGCGGGCTCTCGATATT 7980

7981 TTAAAATGATTGGTCCATGGAGCGCGGGCTCTCGATATTTTAAAATGATTGGTTTGTGAC 8040

8041 GCACAGGCTTTGTTGTGAACCCCATAAAAGCTGTCCCGATTCCGCACTCGGGGCCGCAGT 8100

8101 CCTCTACCCCTGCGTGCTGTACGACTGTGGCCCCAGCGCGCTTGGAATAAAAATCCTCT 8160

8161 TGCTGTTTGCATCAAAAAAAAAAAAAAAAAAAAAA 8195